PROJECT ID # TI095

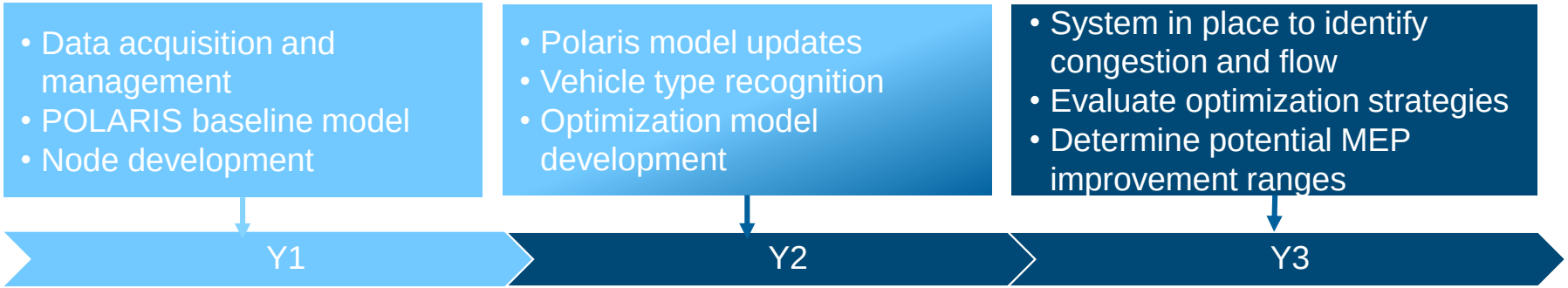
HIGH-PERFORMANCE COMPUTING (HPC) FOR O'HARE HUB

JOSH AULD
Argonne National Laboratory
2020 DOE Vehicle Technologies Office Annual Merit Review
June 2020, Washington D.C.

Project Overview

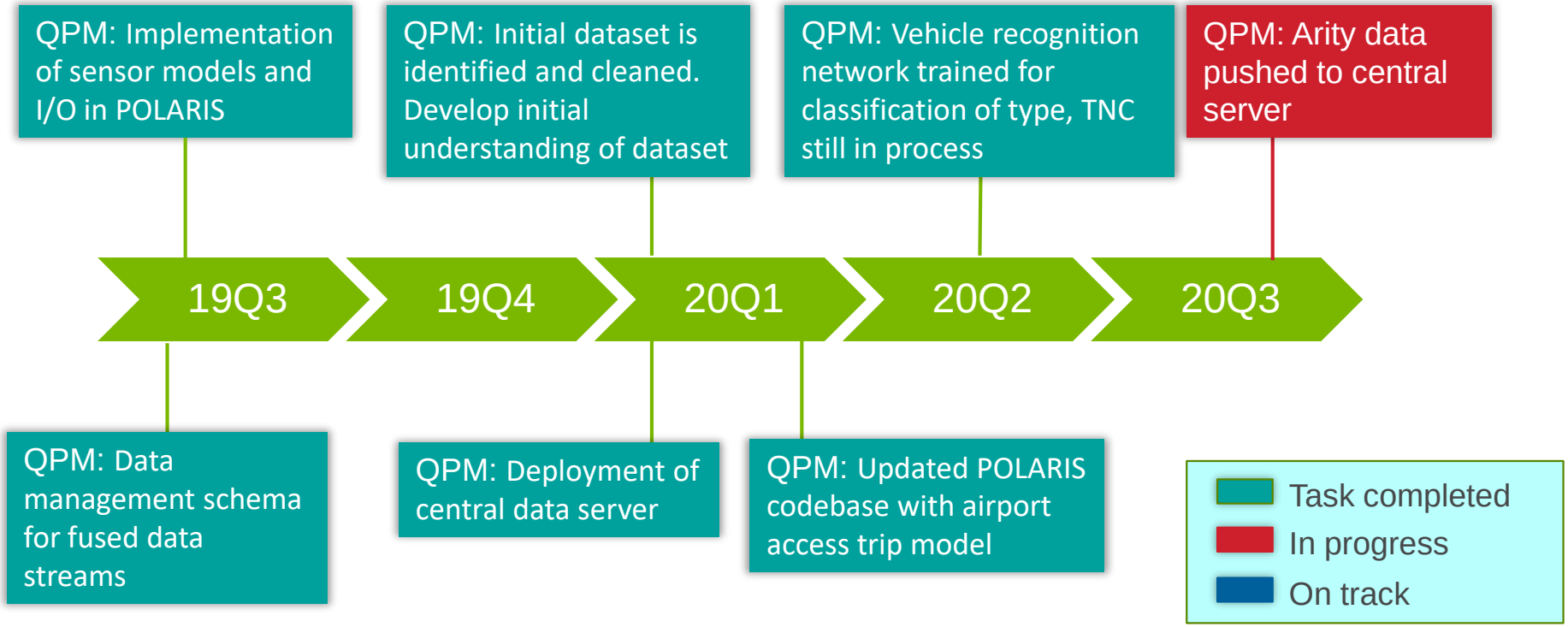
Timeline	Barriers
• Project start date : Jan 2019 • Project end date : Jan 2022 • Percent complete : 35%	• Complexity of large-scale integrated transportation networks • Computational difficulty of accurately modeling and simulating large- scale transportation systems • Accurately measuring the transportation system-wide energy impacts of connected and automated vehicles • Determining the value and productivity derived from new mobility technologies • Difficulty in sourcing empirical real-world data applicable to new mobility technologies such as connectivity and automation
Budget	**Partners**
• FY19 Funding : $1,129K • Total Project : $3,200K	• Chicago Department of Aviation (O'Hare Airport) • Chicago Department of Transportation • Chicago Metropolitan Agency for Planning • Arity • University of Chicago

Project Objectives



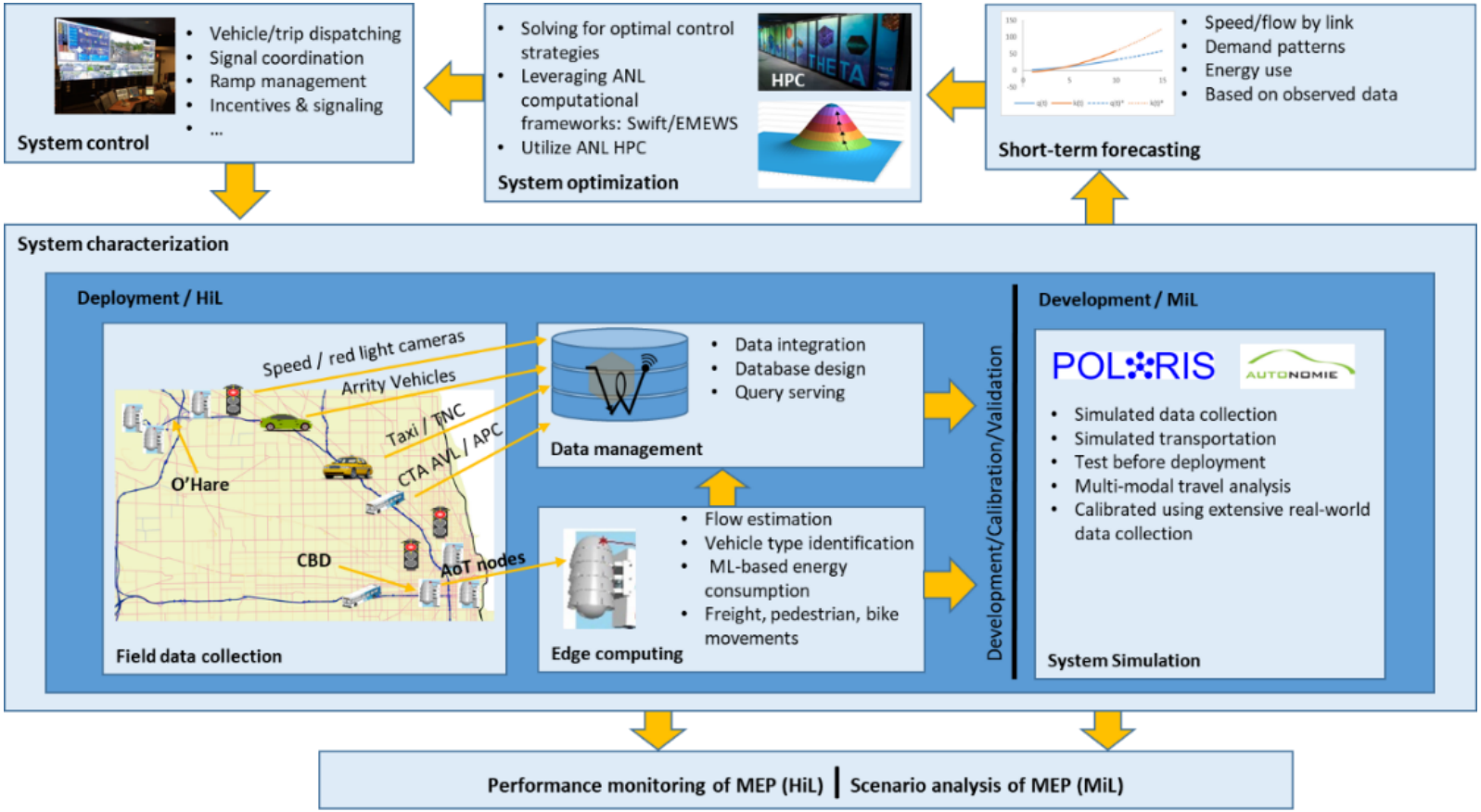
- **VTO Tech Integration goals:**
  - Economic growth: commercialization opportunity for vehicle trajectory data
  - Affordability for business and consumers: increasing transportation efficiency and reducing time cost to access O'Hare transportation hub
  - Reliability/resiliency: increasing O'Hare traffic reliability through data acquisition and active management
- **Impact on addressing barriers:**
  - Add new data sources for real-world transportation system data through Arity and Waggle
  - Accurately model integrated transportation systems using new data and POLARIS
  - Explore system-wide energy impact of active-management using sensing and connectivity

Milestones


QPM: Implementation of sensor models and I/O in POLARIS
QPM: Initial dataset is identified and cleaned. Develop initial understanding of dataset
QPM: Vehicle recognition network trained for classification of type, TNC still in process
QPM: Arity data pushed to central server
19Q3
19Q4
20Q1
20Q2
20Q3
QPM: Data management schema for fused data streams
QPM: Deployment of central data server
QPM: Updated POLARIS codebase with airport access trip model
Task completed
In progress
On track

APPROACH

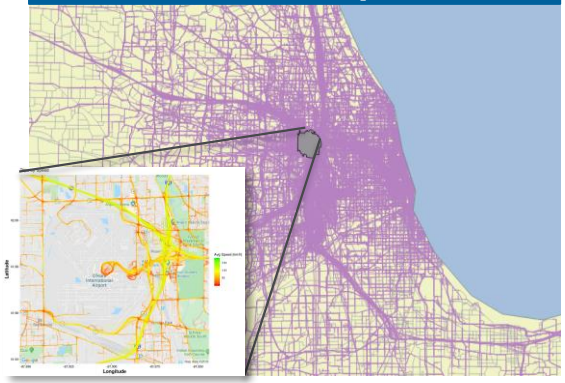
Approach


Vehicle/trip dispatching
Signal coordination
Ramp management
Incentives & signaling
...
System control
Solving for optimal control strategies
Leveraging ANL computational frameworks: Swift/EMEWS
Utilize ANL HPC
HPC
System optimization
Speed/flow by link
Demand patterns
Energy use
Based on observed data
Short-term forecasting
System characterization
Deployment / HiL
Speed / red light cameras
Arrity Vehicles
Taxi / TNC
CTA AVL / APC
O'Hare
CBD
AoT nodes
Field data collection
Data integration
Database design
Query serving
Data management
Flow estimation
Vehicle type identification
ML-based energy consumption
Freight, pedestrian, bike movements
Edge computing
Development/Calibration/Validation
Development / MiL
POLARIS
AUTONOMIE
Simulated data collection
Simulated transportation
Test before deployment
Multi-modal travel analysis
Calibrated using extensive real-world data collection
System Simulation
Performance monitoring of MEP (HiL)
Scenario analysis of MEP (MiL)

Approach: Characterize Traffic State and O/D Flows from Arity and Detector Data

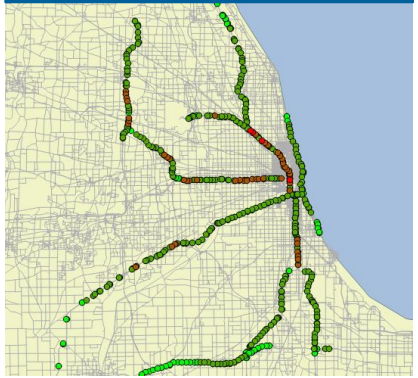
Understand relationship between highway flow and airport demand

Arity: GPS traces for O'Hare trips



- Speed and flow on all traveled links
- Trip origins and destinations
- O'Hare travel routing
- ~1000 trips per day

Live traffic data from IDOT loop detectors



- Continuous speed and flow for highway links at detectors
- Processed ~10 years of data
- Available in near-real time

Short term traffic and demand forecasting

- Predict speed and flow 1-2 hours in advance
- Downstream traffic prediction based on current detector state
- Predict traffic state at O'Hare
- Identify demand patterns (where are trips coming from)
- **Short term state forecasting to be used in optimization**

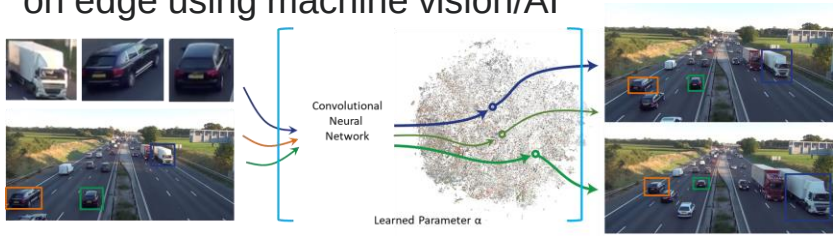
Approach: Enrich the Data with Enhanced Waggle Node Deployment at O'Hare

Identify vehicle types and TNC counts

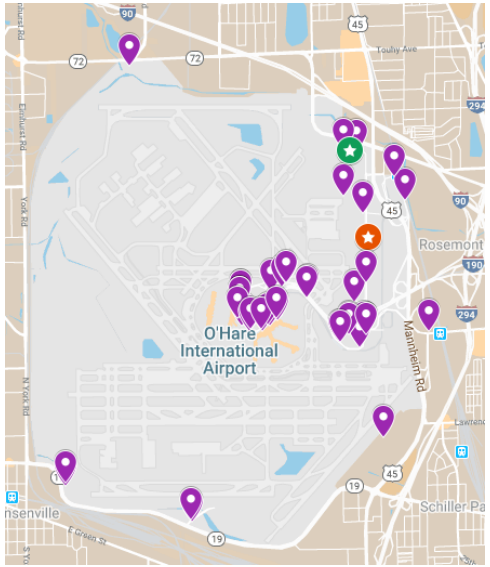
Prototype Waggle Node

Vehicle identification and tracking software running on edge node

Identify vehicle make, model and TNC status on edge using machine vision/AI



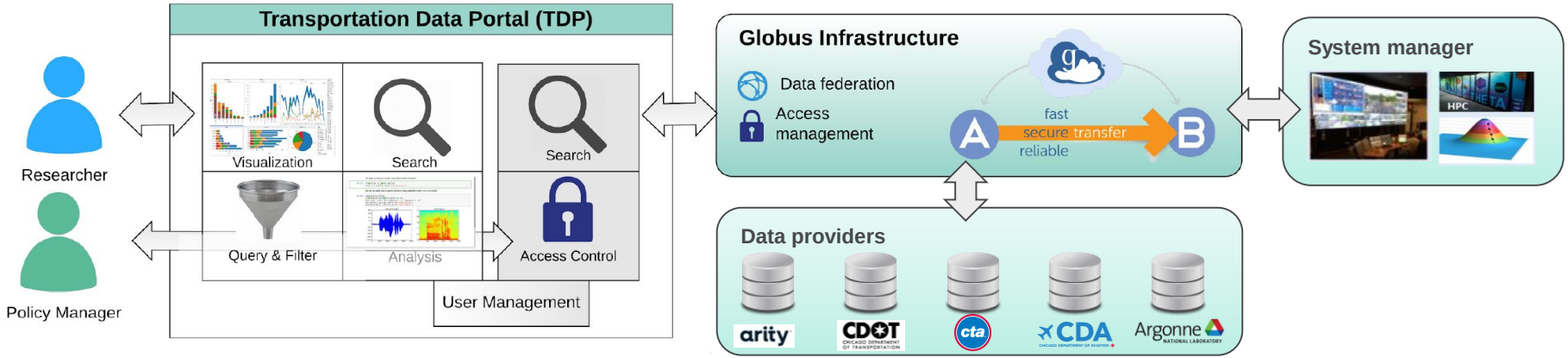
Deploy nodes at O'Hare



Approach: Globus-based Data Management Platform

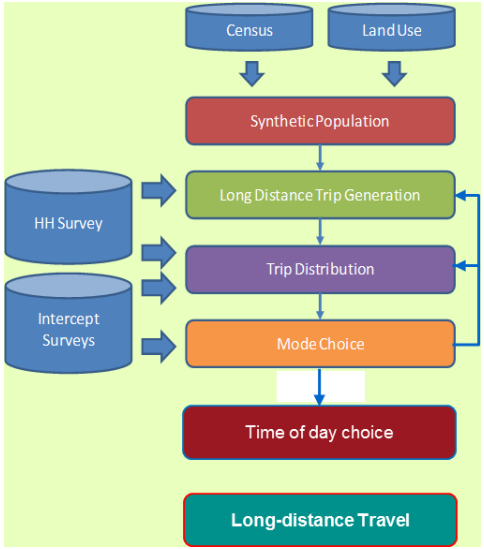
UChicago DOE OS supported tool for data management

- Data management based on Globus platform
  - Open source tool for managing data transfers used in HPC
  - Resources for building research data services – data portal, system manager, etc.
- Allow for connection between disparate data streams, project developed analysis tools and public facing reporting

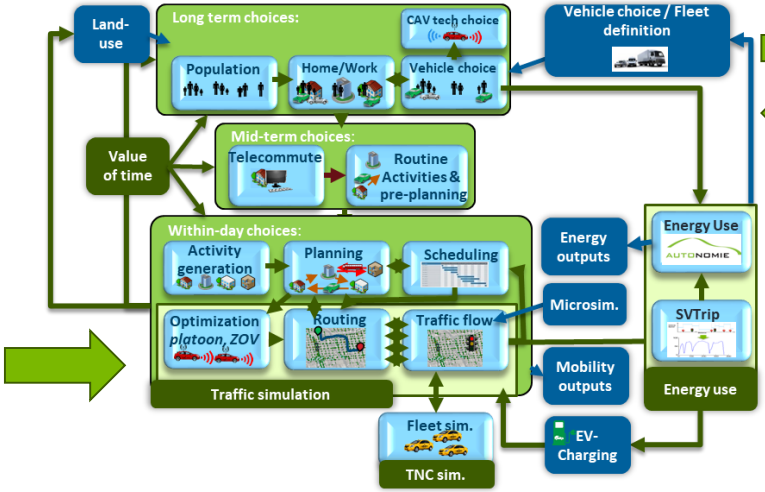


Approach: Develop an Optimization Model for Increasing O'Hare Mobility Energy Productivity

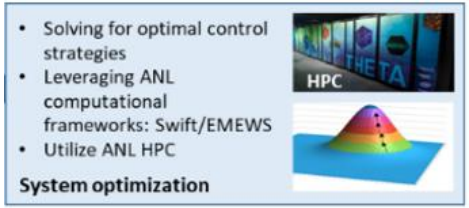
Agent-based Long-distance Travel Model



POLARIS Activity-Based Travel Model



Design and evaluation optimization strategies



Control strategies

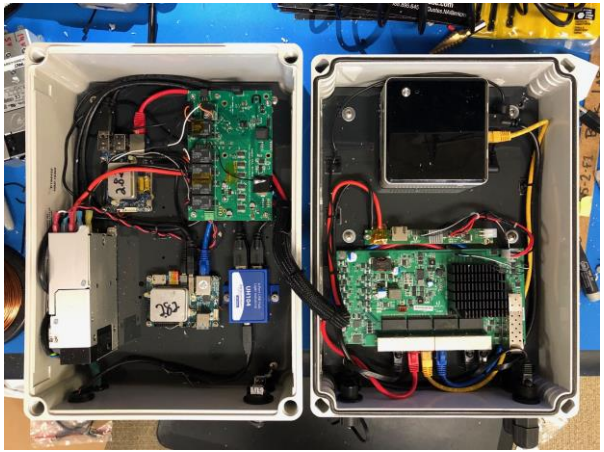
- Dynamic rideshare pricing
- HOT lanes
- Dynamic transit fares
- Dynamic parking charges
- Ramp metering
- Vehicle dispatching

TECHNICAL ACCOMPLISHMENTS

Accomplishment: Node Design Prototyping

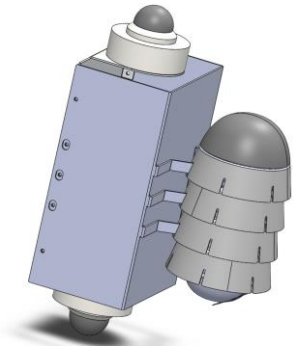
Prototype 1: Currently installed and tested.

Prototype 2: Developed based on lessons learned from *Prototype 1*, with advanced resilience and expandability features. Designed with support for higher quality POE cameras and future edge-class CPUs from NVIDIA, Google and INTEL.



Status:

- Two iterations of design and thermal/indoor testing of *Prototype 2* completed. Outdoor testing soon.
- Delays in production of more prototypes for testing due to supply chain and manufacturing delays caused by COVID-19.



12

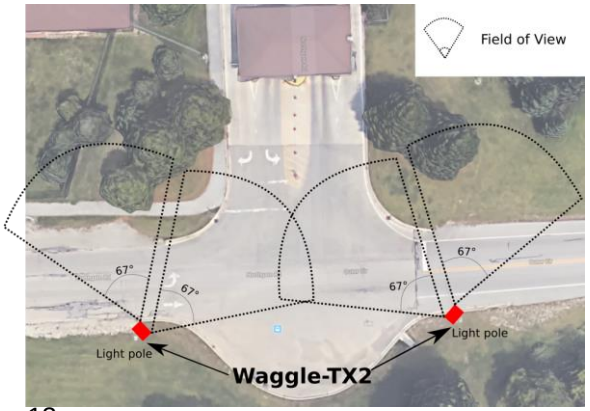
Accomplishment: Testing and Training Data Gathering

- **Two prototype Waggle nodes** equipped with advanced ML/AI computing systems and multiple 4K resolution cameras were **developed and deployed in a test site at Argonne.**




Sample images from the test nodes



Top: Deployed Nodes

Left: Deployment location, orientation and camera viewpoints

Accomplishment: Data Collected from Waggle Nodes

- +22K short (5-30 seconds) video clips have collected from the two prototyped nodes, deployed at the main gate in Argonne campus.
- The clips are used to annotate vehicles and segment them from the background
- Augmenting image collection using online labeling tool:

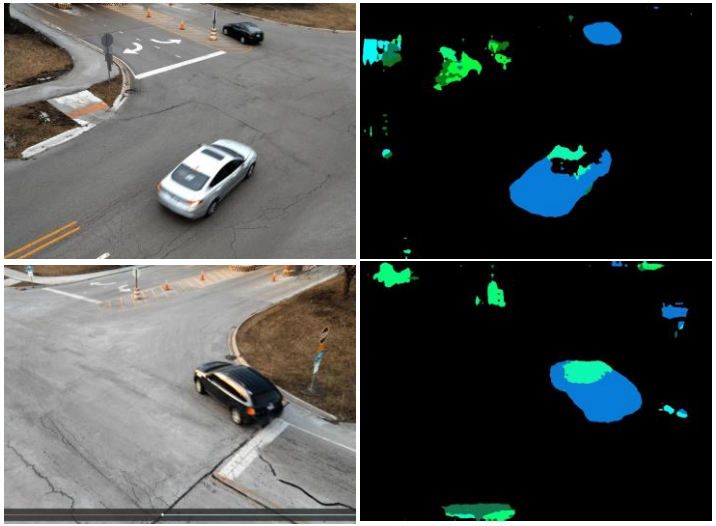
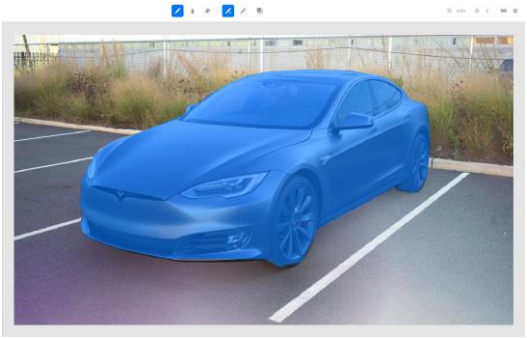


Image segmentation result using FCN-resnet101 on Nvidia TX2; Blue indicates car

LabelBox used to tag images from Waggle with vehicle make/model and TNC status to improve classification models



An example annotation for image segmentation in **LabelBox** (https://labelbox.com/)

Accomplishment: Vehicle Classification and Tracking

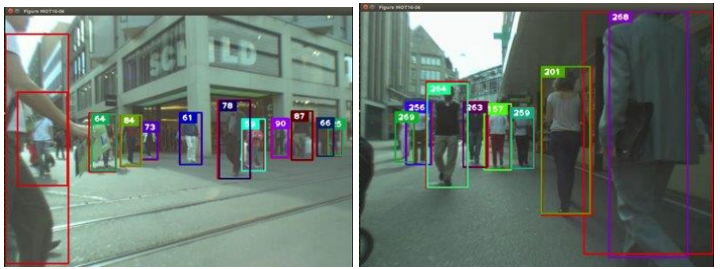
Vehicle make and model classification

- Training a model for multi-perspectives of cars

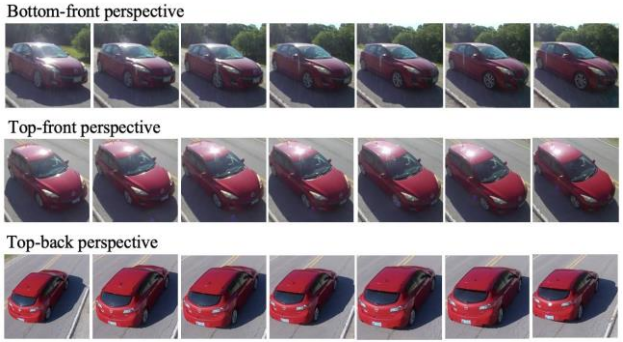
Images from Waggle Deployment on ANL Campus

Object tracker

- Track vehicles/objects for traffic flow estimation

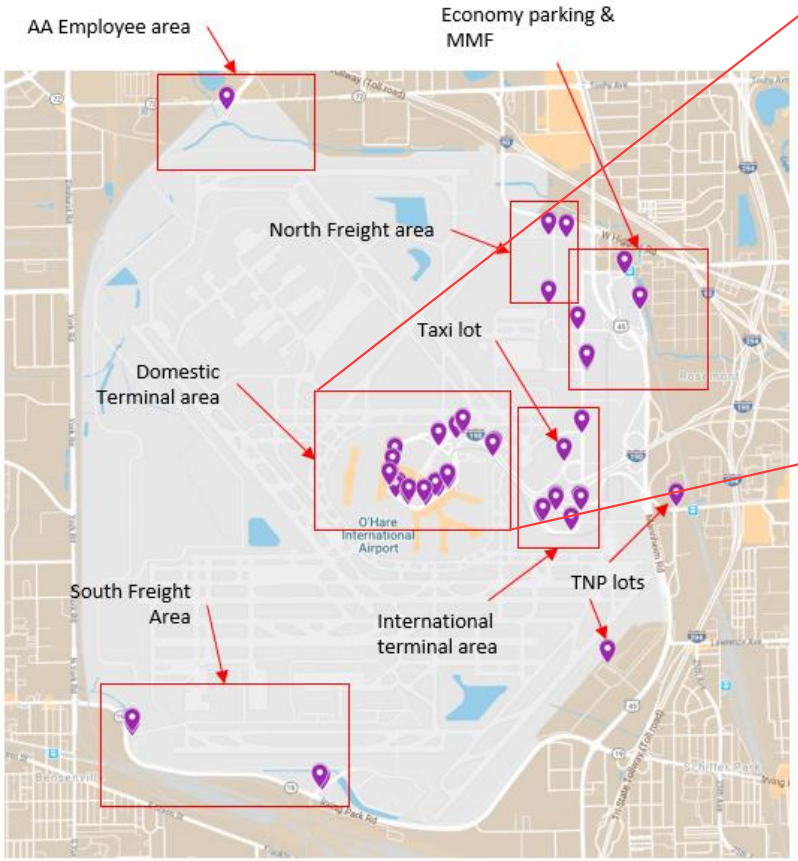


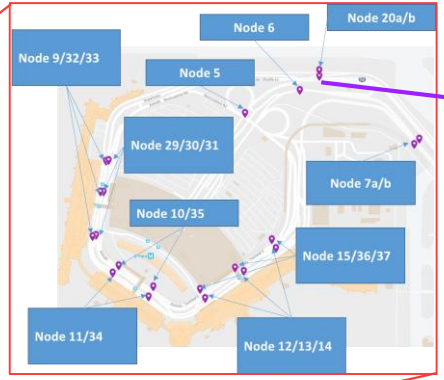
Sample Images from Public Datasets



- Train a model for views from multiple perspectives
- Test configuration:
  - PyTorch on Jetson TX2
  - YOLO v3
  - ANL Main gate dataset
  - Overall accuracy: 50% (but with over 90% confidence)

Accomplishment: Node Locations Identified

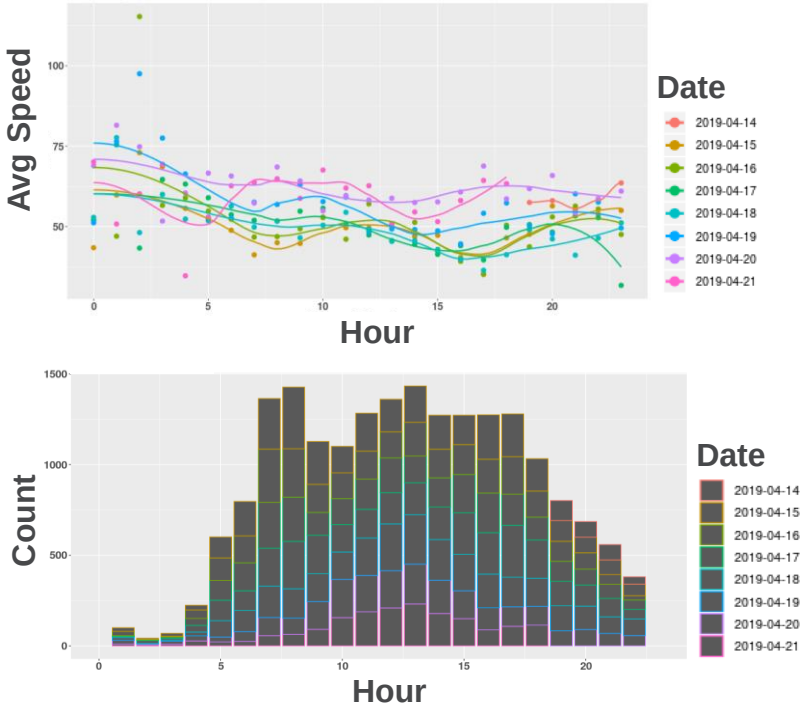




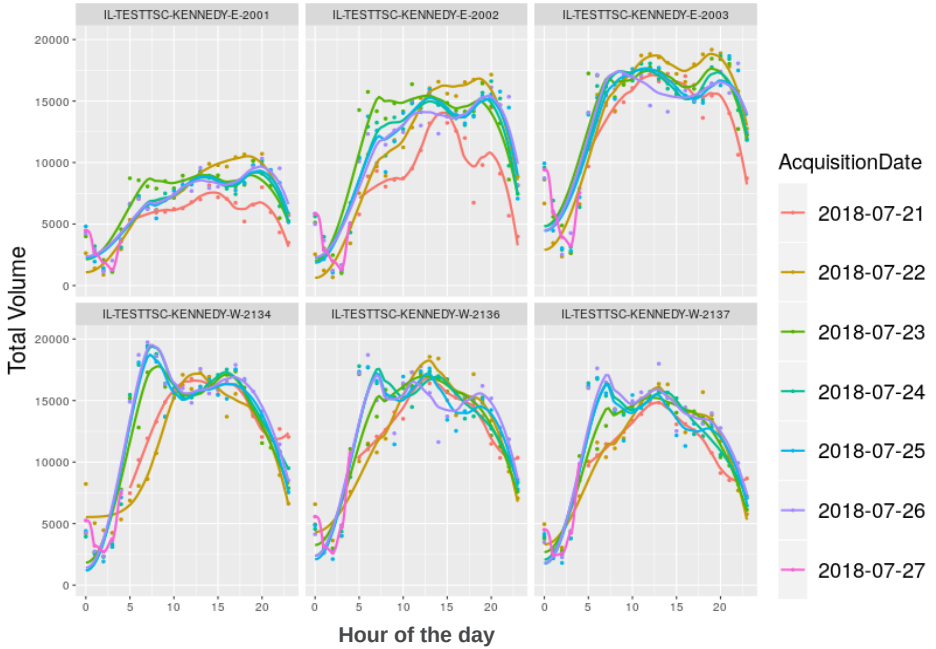
- Multiple site visits with O'Hare ground operations team
- Site survey to verify power availability and cellular connectivity at each node location
- O'Hare approval for node deployment
- Working with pre-approved electrical contractor for installation

Accomplishment: Analyzed Arity and Loop Detector Data

Arity speed and volume by time of day



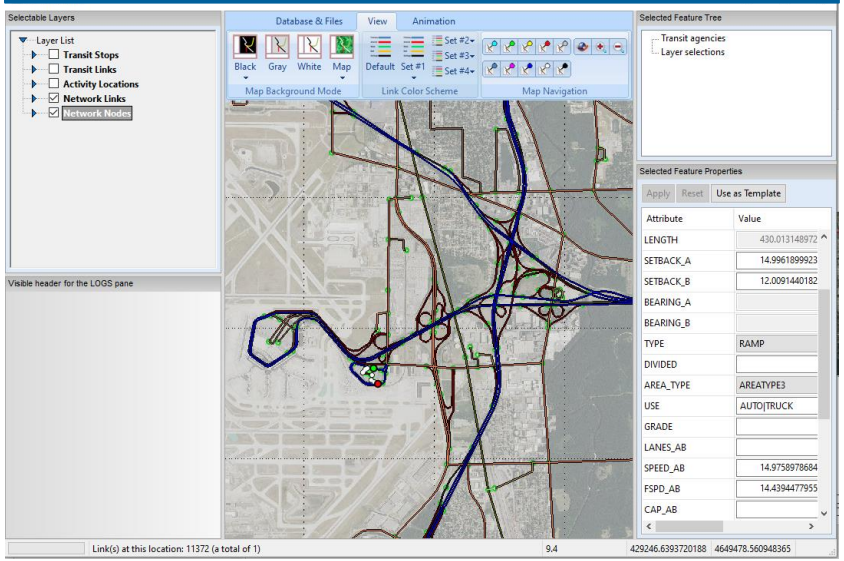
Matched to IDOT loop detectors



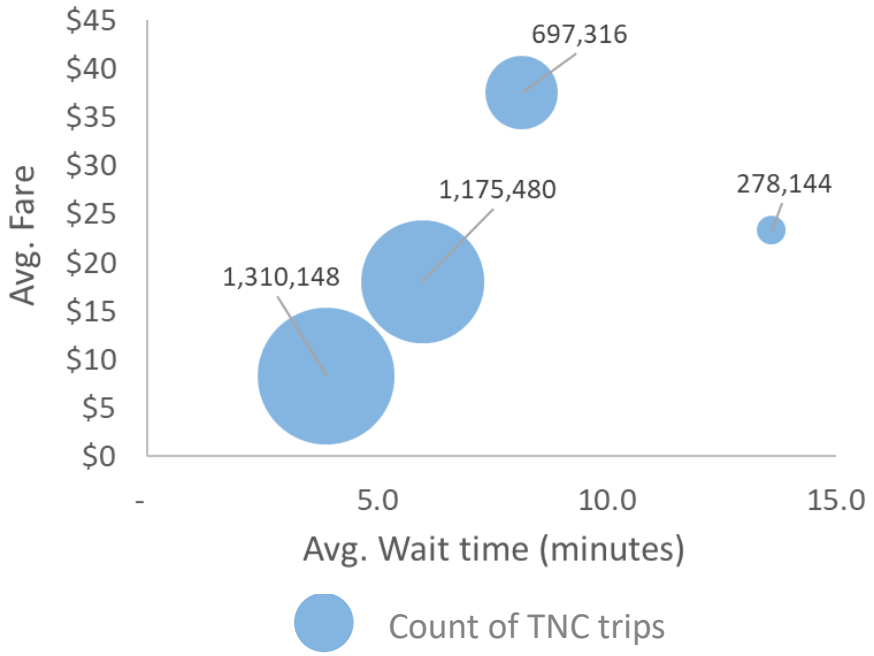
Developing relationships between Arity and IDOT observations for short-term forecasting

Accomplishment: POLARIS Model Updates

Updating physical model of O'Hare area



TNC cost sensitivity study in POLARIS



Collaborations and Coordination

arity	Vehicle trajectory data
CHICAGO CDA DEPARTMENT OF AVIATION	Stakeholder, site access, ground operations data
CDOT CHICAGO DEPARTMENT OF TRANSPORTATION	Stakeholder, data provider
THE UNIVERSITY OF CHICAGO	Globus platform and endpoint for data management
Chicago Metropolitan Agency for Planning	Stakeholder, long term planning support
	TNC fleet modeling

Overall Market Impact: Sustainability and Replication

- Developing traffic and vehicle energy sensing platform
  - Identify make and model of cars to allow for energy estimation
  - Traffic flow characteristics
  - Characterizing traffic conditions and OD flows at key transportation hub
- Based on existing Waggle platform that is extensively deployed
  - Allow for quick expansion to other key transportation hubs
  - Expanding capability of Waggle nodes will allow for project sustainability, especially if already deployed nodes can be upgraded
- Expanding sensing and coupling with other data sources in a management platform will allow for development of traffic management and optimization capabilities

Overall Market Impact: Future Work

- Deployment of Waggle nodes at selected O'Hare sites
- Implementation of ML modeling for TNC vehicle identification
- Field testing and verification of vehicle make/model identification
- Development of short-term traffic forecasting and demand forecasting models
- Expanded case-studies in POLARIS evaluating sensitivity to key optimization parameters

Any proposed future work is subject to change based on funding levels

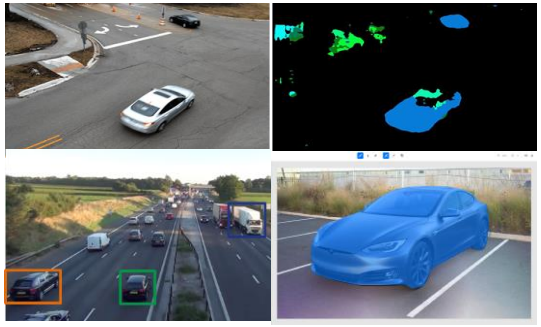
Overall Market Impact: Remaining Challenges and Barriers

- Node deployment challenges due to COVID19
  - Manufacturing delay for new Waggle nodes
  - Installment delay – working with electrical contractor for deployment
- Reduction in air travel due to COVID19 limits representativeness of data
- Camera resolution on Waggle nodes
  - Testing higher resolution cameras
  - Requires substantial on-board processing
- Real-time data access
  - Delays in data provided by IDOT for highway sensor
  - Computational burden in real-time processing on data management platform

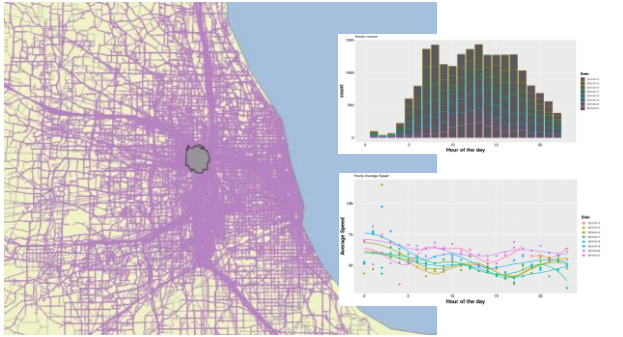
Summary

Node Development

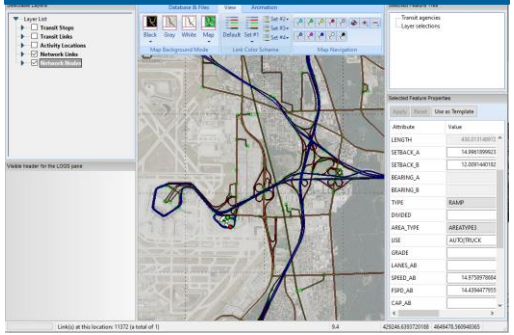
Vehicle Classification



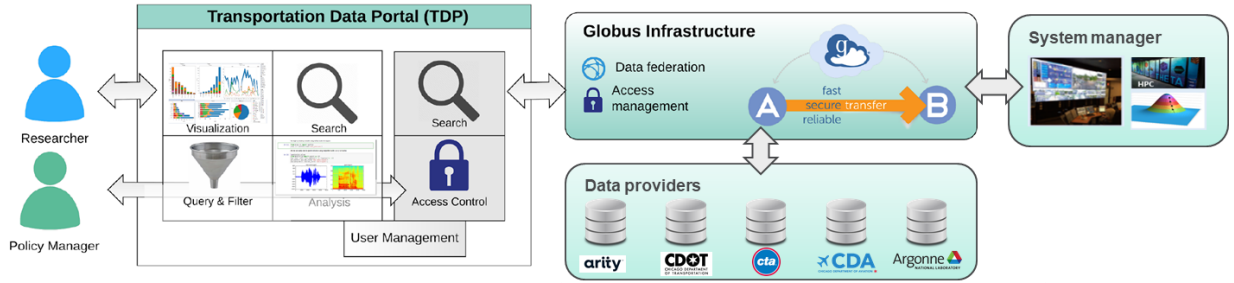
Data Analysis



POLARIS Modeling



Data Management Platform



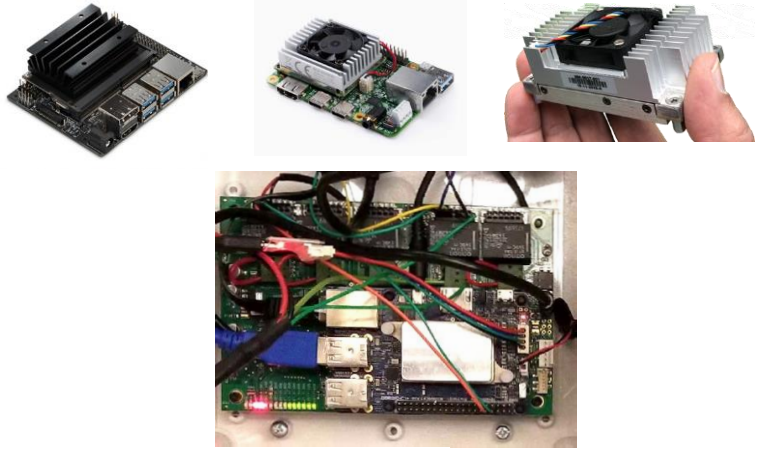
TECHNICAL BACK-UP SLIDES

Node Design: Software and Hardware

Powerful edge devices for AI@Edge

- Hardware support including Tensorflow, OpenCV/CL, PyTorch APIs through NVIDIA and Google ML/AI focused devices in Waggle. Initial nodes will have TX2, and the future nodes may include TX2/Xavier NX along with Google Edge TPU.

- Software support through docker containers. Enhanced node OS to support deployment of all applications through docker and Kubernetes based framework.



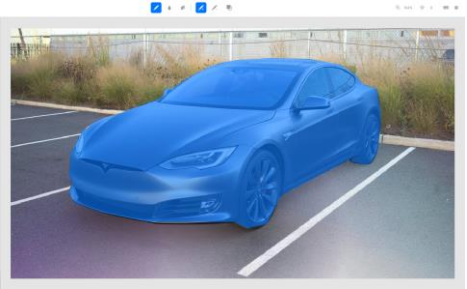
Hardware integration

Software integration

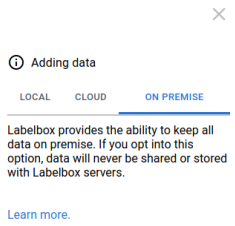
Annotation of Collected Training Images

- To accelerate annotation on the large set of image/video training samples, we are evaluating online image annotation tools that can support crowd sourcing of labels
- Features being considered -
  - Fast annotation: Each vehicle annotation in about 30 seconds by a human
  - Review mode: allow people to verify the annotations by others and vote/enhance.
  - Consensus: Tools also allow consensus between multiple annotations
  - Ability to support fanning out jobs to large number of ANL employees who are unable to perform
  - their regular job duties due to mandatory COVID-19 related lab closure.

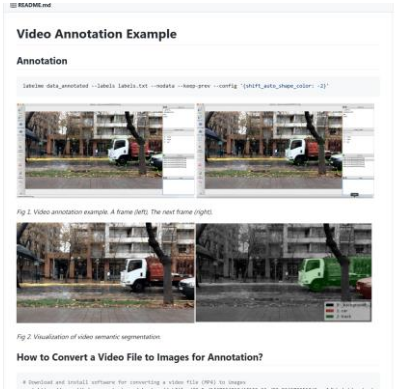
Considering 2-3 COTS and open-source tools ...



An example annotation for image segmentation in **LabelBox** (https://labelbox.com/)



On premise option allows working with images without uploading them to external server

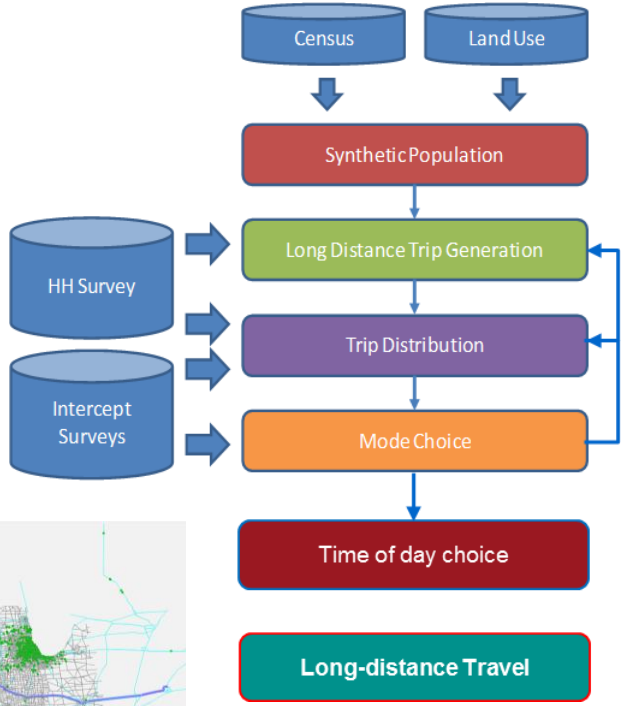


Labelme open-source tool provides similar features (https://github.com/wkentaro/labelme/tree/master/examples/video_annotation)

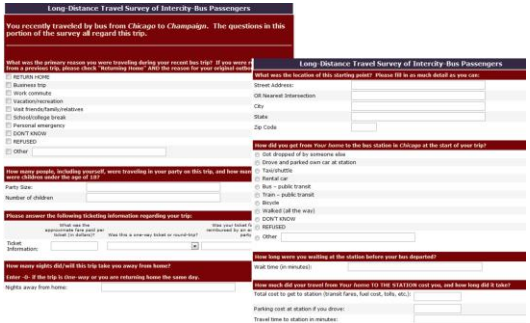
Illinois Long-Distance Travel Model Framework

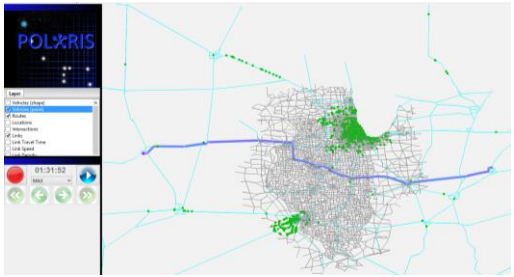
Includes visitor and resident airport travel model

- Primary inputs:
  - Census data (ACS and 2010 SF1)
  - TAZ Land use data from MPOs
  - Congested Network skims
  - **Person and intercept survey** results
- Five inter-related models
  - Generation, Distribution, Mode choice connected through logsums.
  - Conditional time-of-day choice
  - Population synthesis using PopSyn program developed for CMAP



Long-distance intercept survey conducted in 2012





Argonne NATIONAL LABORATORY

REVIEWER ONLY

Critical Assumptions and Issues

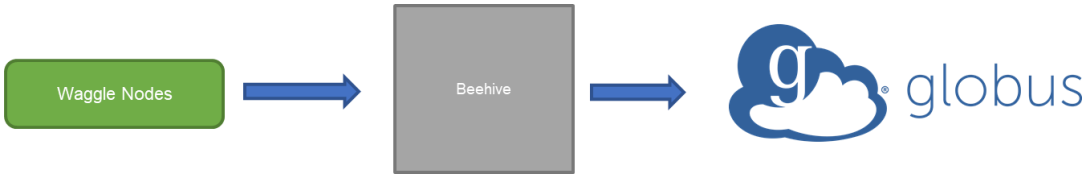
- Delay in deployment of new Waggle nodes at O'Hare site due to supply chain and manufacturing issues related to COVID-19
  - Waiting for manufacturer to restart
  - Requires stay-at-home order to be lifted in order to have contractor start installing nodes once received
- Massive change in air-travel patterns (~95% reduction in air-travel) makes data over current period non-representative for long term planning
  - Re-order project timeline to focus on simulation and machine learning improvement while waiting for air travel to recover and new nodes to be deployed

Publications

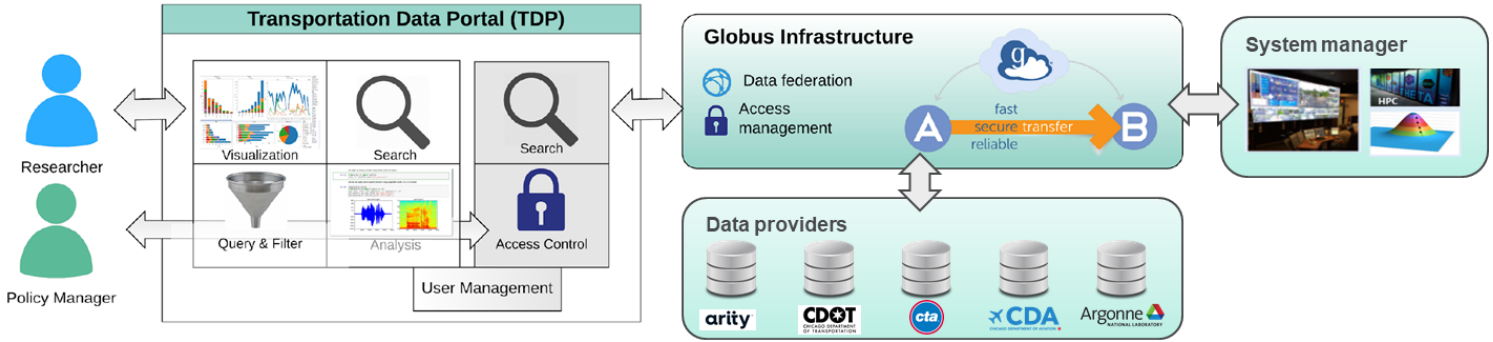
- Papers will be presented upon completion of the project

Data Management Plan

- We have set up a Globus entry point in the server
- All data has secured 2-factor protected access
- Image/video samples collected in the prototype test nodes are first staged in Beehive, the cloud server managed by the Waggle team, and next synchronized with the Globus entry point.
- Data from different sources will be merged through Globus and used in the analysis tasks of the project.



Globus-based data management platform



Past Comments From Reviewers

- This project was not reviewed last year.